

A new consideration on the formation of sunspots

Abstract

It is accepted that sunspots occur when a magnetic flux tube rises above the photosphere, forming a pair of positive and negative spots at the cross-sections. However, this model has at least three long-standing problems, which contradict with the model, but have been left behind for a long time. They are: (1) the presence of magnetic monopole-like single spots, (2) the delay of the appearance of *f*-spots after *p*-spots, (3) the flux unbalance between *p*- and *f*-spots. Further, there are two newer observations to solve. They are: (4) positive single spots tend to be present in positive unipolar magnetic regions (vice versa); (5) pairs of spots tend to be present at the boundary unipolar magnetic regions, not elsewhere. The aim of this paper is to suggest that the five problems (1–5) may be explainable all together by considering: (a) the unipolar (positive and negative) magnetic regions are one of the basic structures generated by the internal solar dynamo, not just remnants of old active regions; (b) single positive spots (unipolar magnetic spots) are the basic unit of sunspots and are born in a positive unipolar magnetic region (vice versa), and (c) a *p*-spot induces a *f*-spot across the boundary of two unipolar magnetic regions, forming pairs of spots (positive *p*-spots in positive unipolar regions and negative *f*-spots in the negative unipolar regions).

Keywords: magnetic flux tube, photosphere, unipolar magnetic regions, single positive spots, pairs of spots

Introduction

The formation of sunspots has been considered by many researchers in the past. After the discovery of magnetic fields in sunspots by George E. Hale, magnetic fields of sunspots have become one of the main issues of sunspots. Early studies of the formation of sunspots are summarized by Abetti,¹ Cowling², Kiepenheuer³, de Jager⁴, Menzel⁵, Bray and Loughhead⁶. A recent comprehensive review on sunspots was made by Solanki⁷.

Among a few proposed theories, the rising magnetic flux tube model by Babcock⁸ has long been accepted as a unique model in considering the formation of sunspots (hereafter, this model is referred as the RF model). In the model, sunspots appear as a pair of positive and negative spots as a magnetic flux tube breaks through the photospheric surface at two points. Thus, a pair of positive and negative forms the basic unit. In spite of a great observational and theoretical progress in studying sunspots during the last few decades, these studies have essentially based on the RF model.

Thus, a few long-standing problems, which are contradictory with the RF model, have remained unsolved until now. They are:

- i. The presence of magnetic monopole-like single spots.
- ii. The delay of appearance of *f*-spot after *p*-spot.
- iii. The unbalance of the magnetic flux between a *p*-spot and a *f*-spot.

In addition, there are two more recent observations, which are not easily be explained (McIntosh⁹; see also Akasofu^{10,11}): They are:

- iv. Positive single spots tend to appear in a positive unipolar region (vice versa).
- v. Pairs of spots tend to appear at or near the boundary of two unipolar magnetic regions, not in the middle of unipolar region.

Fortunately, the new efforts of attempting to understanding the observations (4, 5) have provided important hints to study the above three problems (1, 2, 3).

Volume 2 Issue 5 - 2018

Syun-Ichi Akasofu

International Arctic Research Center, University of Alaska Fairbanks, Alaska, USA

Correspondence: Syun-Ichi Akasofu, International Arctic Research Center, University of Alaska Fairbanks, PO Box-757340 930, Koyukuk Dr. 415, Akasofu bldg Fairbanks, Alaska 99775 7340, USA, Tel 9074 7460 12, Email sakasofu@alaska.edu

Received: August 31, 2018 | **Published:** September 17, 2018

The purpose of this paper is to suggest that the five problems (1–5) may be explainable all together by considering: (a) unipolar magnetic fields are one of the basic features in solar magnetism related to the solar internal dynamo, not just remnants of old active regions, (b) single spots are the basic unit of sunspots, not pairs of spots. (c) a spot induces a *f*-spot across the boundary of two unipolar magnetic fields (of the opposite polarity). This work is mostly based on the magnetic data based on the Kitt Peak Solar observatory.

Unipolar magnetic regions

Morphology

Before examining the formation of sunspots, it is important to remind about the distribution of magnetic fields on the photospheric surface. Figure 1 shows an example. First of all, there are weak, but large-scale positive and negative fields. They have been called “large-scale fields”,⁹ and are unipolar magnetic regions. Secondly, there are much smaller, but concentrated fields. They are pores and small sunspots; note that positive ones are in positive unipolar regions (vice versa). Thirdly, some of these small positive and negative fields are concentrated together near the boundary of two unipolar regions (of the opposite polarity), again positive ones are located in a positive unipolar region (vice versa).

Figure 2 shows a typical solar magnetic record in the latitude-longitude map. The region marked “1” are positive (red) and negative (blue) unipolar magnetic regions, respectively. The regions marked by “2” have pores, single spots and groups of single spots (positive and negative). A distinct group of pairs of positive and negative spots near the boundary of two unipolar magnetic regions are marked by “3”. As can be seen in later examples, this general pattern is maintained throughout the sunspot cycle, although it becomes more complex when the sun becomes active.

In Figure 2, it may be recognized: (a) groups of positive pores, single spots and groups of spots tend to be present in a positive unipolar regions, and vice versa, (b) pairs of spots tend to be present at or near one of the boundary of two unipolar regions, not in the middle of unipolar regions, and further positive spots are in a

positive unipolar region (vice versa). These features can also be seen throughout sunspot cycles; Figures 5A–5E.

On the basis of Figure 1 & Figure 2, and a large number of others, including those which will be seen later. Figure 3A shows schematically the basic distribution of the unipolar regions. Positive and negative unipolar regions alternately distributed in longitude.

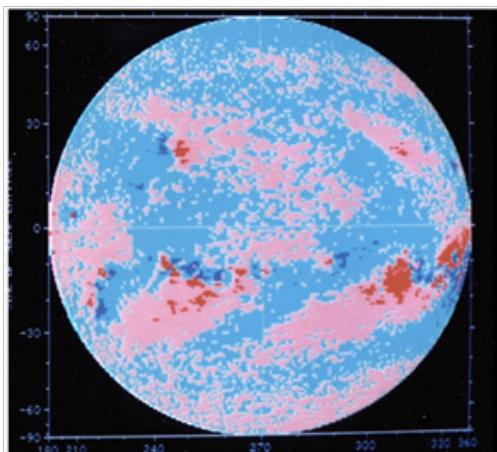


Figure 1 A typical example of the distribution of magnetic field seen on the disk of the sun (The Kitt Peak Observatory).

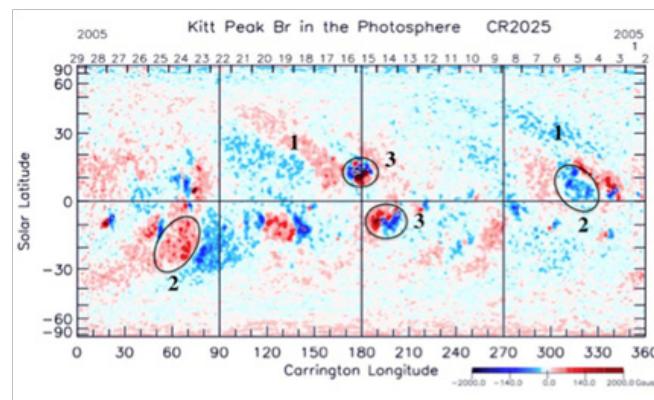


Figure 2 The distribution of the photospheric magnetic field map (longitude-latitude) during the solar rotation CR 2025. The regions marked by "1" indicate positive (red) and negative (blue) unipolar magnetic regions, respectively. The regions marked by "2" indicate pores, single spots, groups of single spots of positive and negative polarities. The region marked by "3" is the location of pairs of positive and negative pores and spots. There are also unipolar regions in the northern southern highest regions, the polar unipolar magnetic regions (The Kitt Peak Solar Observatory).

In this study, it is crucial to have a continuous color presentation of the magnetic fields, because all the above features have a wide and continuous range of the intensity (0–2000G) from very weak unipolar regions to most intense sunspots; The recent black–white presentation with a high resolution tends to miss the important features, which are essential in this paper; in particular, it is not easy to see the crucial relationship between the unipolar magnetic regions (which tend to have a low intensity) and sunspots; see Appendix.

Unipolar magnetic regions

Morphology

It has long been considered that unipolar magnetic regions are

remnants of old active regions (Figure 3B) on the basis of a study of the sunspot cycle by Leighton¹², in which active regions are elongated by the non-uniform rotation of the sun. In this section, it will be shown that unipolar regions are different from old active regions and thus they should be distinguished. Further, it is suggested that unipolar regions are one of the fundamental features of solar magnetism, which have the origin in the internal dynamo.

(a) Latitudinal extent

One of the important features which distinguish between unipolar regions and old active regions is their difference in the latitudinal extent. Unipolar regions extend from 10° to more than 60°, and often connect to the polar unipolar regions and also connect to unipolar regions of the same polarity across the equator, while the sunspot regions extend from 0° to at most 45° as the Butterfly diagram shows. Figure 4A shows an example of their connection to both the polar and equatorial regions. This polar connection may be related to the poleward migration of unipolar magnetic regions during a certain periods¹³ and thus related to the reversal of the dipolar field of the sun and thus to the internal solar dynamo¹⁴ (Figure 4B).

(b) Development

However, an even more important difference between them is the fact that unipolar regions appear before sunspots are born in them at the early epoch of the sunspot cycle. It is unfortunate that such a fundamental importance of unipolar regions and their importance on the formation of sunspots have not been recognized in the past.

Figure 5A represents the solar magnetic field during the earliest epoch and early epoch of a new sunspot cycle. It is particularly important to note that unipolar regions just begin to appear before any sunspots appear. This indicates clearly that unipolar magnetic regions are not remnants of old active regions. This fact becomes more clear in Figure 5B; it can be seen that a pair of unipolar magnetic regions appeared most clearly without sunspots, indicating the growth of unipolar regions without sunspots and thus before the birth of spots. Thus, they cannot be remnants of old active regions.

Figure 5C shows two magnetic records of the rising and maximum epochs of the cycle. One can see that unipolar regions grow in intensity and latitudinal range (toward both higher and lower latitudes) as the cycle advances. Sunspot activities tend to distort the basic pattern of unipolar magnetic regions because of sunspot motions as the sunspot cycles advances. Figure 5D shows two records during the declining epoch of the cycle. As sunspot activities subside, the basic pattern begins to be restored.

Figure 5E shows two examples of magnetic records in the final epoch of the cycle. It is important to recognize that the unipolar regions tend to fade away in high latitudes, but remain only in the lowest latitude. Thus, a new cycle (Figures 5A & 5B) begins in high latitudes without remnants of old active regions.

Thus, in summary, as pointed out by McIntosh & Wilson¹⁵ and Wilson¹⁶, the basic unipolar magnetic regions are different from remnants of old active regions, so that unipolar magnetic regions should be distinguished from old remnants of active regions.

Based on a large number of records during more than two sunspot cycles, Figure 5F summarizes schematically the close relationship among single spots, pairs of spots and unipolar regions, including also the well-known other features of pairs of spots, such as the reversal

of the polarity of spots in the northern and southern hemispheres and during the even and odd sunspot cycles, and also the p -spots are located a little equatorward of f -spots.¹¹ In the following, it will be shown that unipolar magnetic regions have a crucial importance in explaining the five problems (1–5).

Model

Based on the observed facts in this section, it can be said that the unipolar magnetic regions are important in reconsidering morphology of sunspots and that the unipolar regions seem to have the origin in the internal dynamo. Thus, it is worthwhile to refer here to the internal dynamo models of solar magnetism. Among many internal dynamo models, the theoretical model developed by Nakagawa¹⁷ considers specifically the origin of positive and negative unipolar regions as shown in Figure 3C: he noted: “The unipolar regions in Figure 1 are identified with the variations of the magnetic field near the surface, so that the flux coil which appeared near the surface is identified as one polarity, while the submerged region is the opposite polarity.”

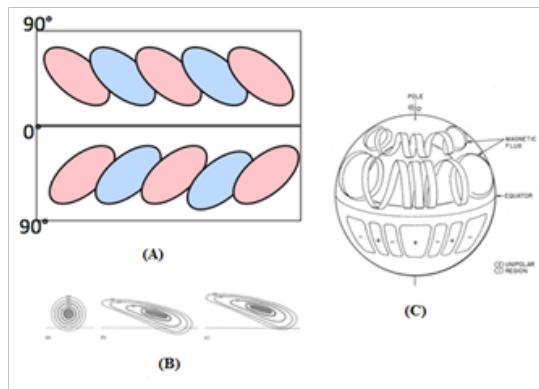


Figure 3 (A) Schematic illustration of the positive and negative unipolar regions in both hemispheres. (B) The development of an old active region.¹² (C) The solar dynamo model developed by Nakagawa¹⁷. It shows magnetic fluxes in the northern hemisphere. Its photospheric fields are shown as unipolar magnetic field regions in the southern hemisphere.

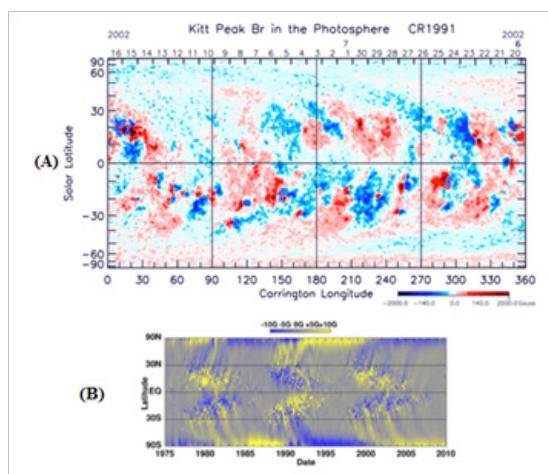


Figure 4 (A) Unipolar regions are often connected to both the polar unipolar regions and also across the equator. (B) The migration of middle latitude magnetic fields to the polar unipolar magnetic regions.¹³

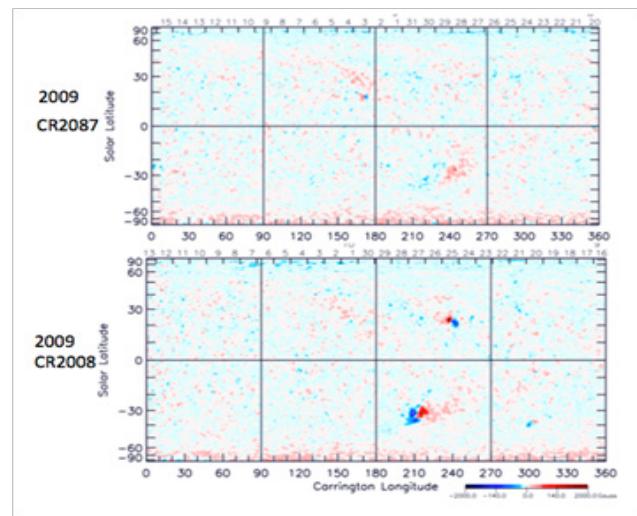


Figure 5A Two solar magnetic records during the earliest epoch of the solar cycle. They show the birth of newly developed unipolar regions.

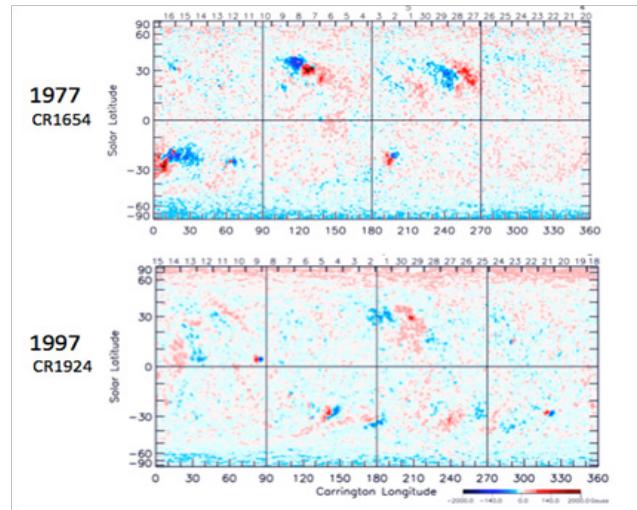


Figure 5B Two magnetic records during an early epoch of the cycle. Compared with Figure 5A, one can see clearly the development of unipolar regions.

Summary

In summary, the fact that the latitudinal extent and growth during the sunspot cycle, together with the fact that unipolar regions can be connected to the polar unipolar regions, as well as Nakagawa's internal dynamo model, suggest that *unipolar magnetic regions are one of the basic features in solar magnetism related to the internal dynamo*. Thus, unipolar magnetic regions are different from remnants of old active regions,^{15,16} so that they should be distinguished from old remnants of active regions.

Further, as shown in the next section, positive single spots appear in a positive unipolar region (vice versa), so that positive unipolar regions are the birth place of positive single spots (vice versa). This interpretation of unipolar magnetic regions is crucial in explaining the problems (1–5) as mentioned earlier.

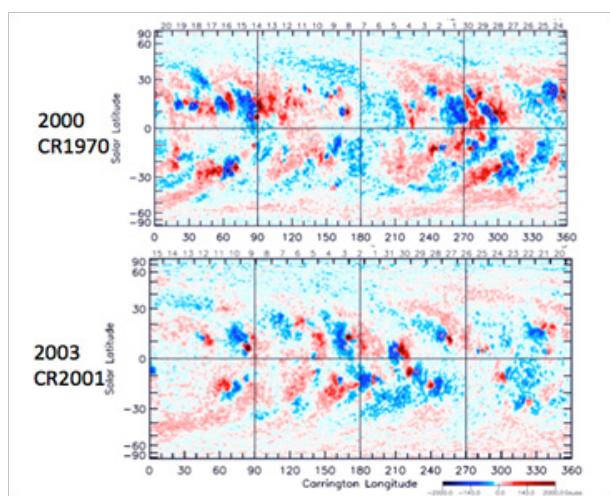


Figure 5C Two magnetic records during the rising and maximum epochs of the cycle.

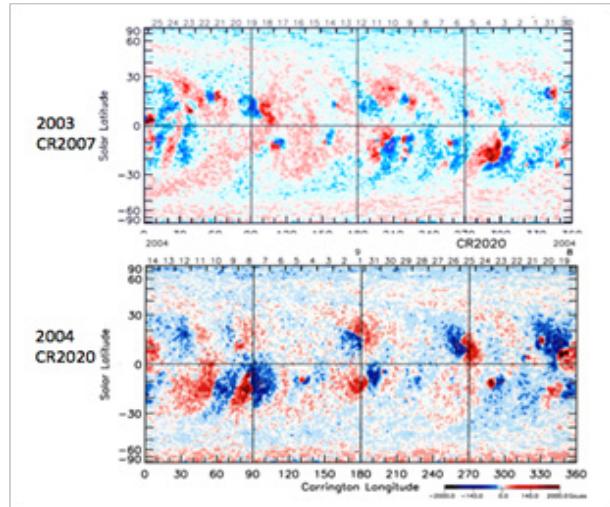


Figure 5D Two magnetic records of the declining epoch of the cycle. The general features pointed out in Figure 1 can be recognized.

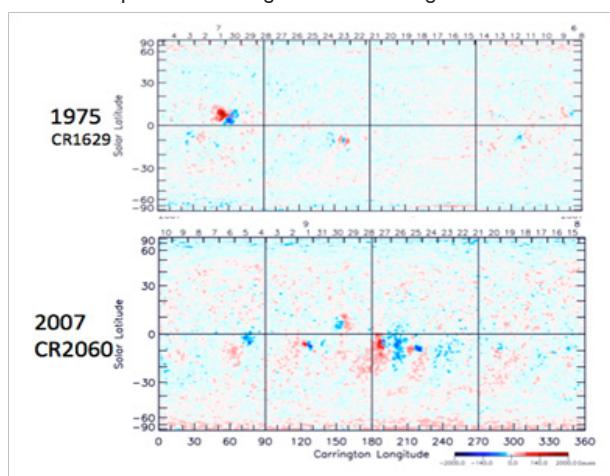


Figure 5E Two magnetic records during a very late epoch of the cycle. It may be noted that compared with Figures 5C and 5D, unipolar regions became very weak and are almost faded away.

Single spots

Pores and single spots

Although single magnetic monopole-like spots (or often called ‘unipolar’, ‘isolated’ or ‘independent’ spots) are illustrated in many books or review papers,^{1,6,7,18} they have not been considered in recent years.

Figure 6 shows two examples of a single spot. Since sunspots must appear as a pair of positive and negative spots in the RF model, the presence of single, isolated or independent spots has long been a puzzle, in addition to the fact that magnetic mono-poles cannot exist. Single spots we study here should not be confused with some of scattered single spots in old active regions during a late phase of active sunspot groups, the $\beta\gamma$ classification in Bray & Loughhead⁶; they will be referred later.

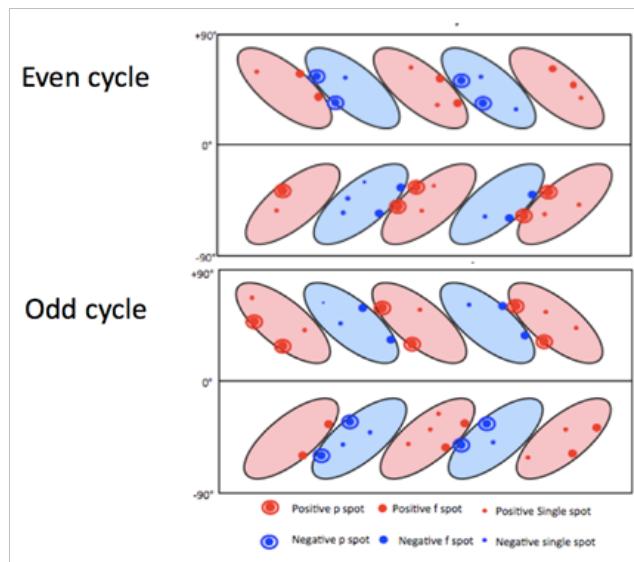


Figure 5F Assembly of the available information regarding single spots, sunspot pairs and their relationship among positive and negative unipolar regions, for both even and odd cycles.¹¹ Other known facts, such as the reversal of the polarity with respect the equator and sunspot cycle are also included.

Since magnetic monopoles cannot exist, we try to consider the formation of single pores and spots as a *local* phenomenon within unipolar magnetic regions, namely *a positive single spot occurs in a positive unipolar region (vice versa)*. The clue for this consideration is the observation (4) that positive single spots appear in a positive unipolar region (vice versa).

Figures 7A & 7B show two high resolution images of a single spot. They show that a single spot consists of several pores or small spots,¹⁹ which indicate that a single spot is an assembly of several small spots, not a simple column of magnetic flux. There is a possibility that $\beta\gamma$ spots may be disassembled small spots, rather than disintegration of a single column of the flux tube.

Before examining the formation of sunspots in unipolar magnetic regions, we examine unipolar magnetic regions with a high resolution image. Figure 8 shows an example, which may be compared with Figure 1. It can be seen that the high resolution image of unipolar magnetic regions shows a magnetic network structure. It has been

considered that the network structure is formed by the convective flow in each cell, in which flows continuously sweep magnetic fields to its boundary and send them down.

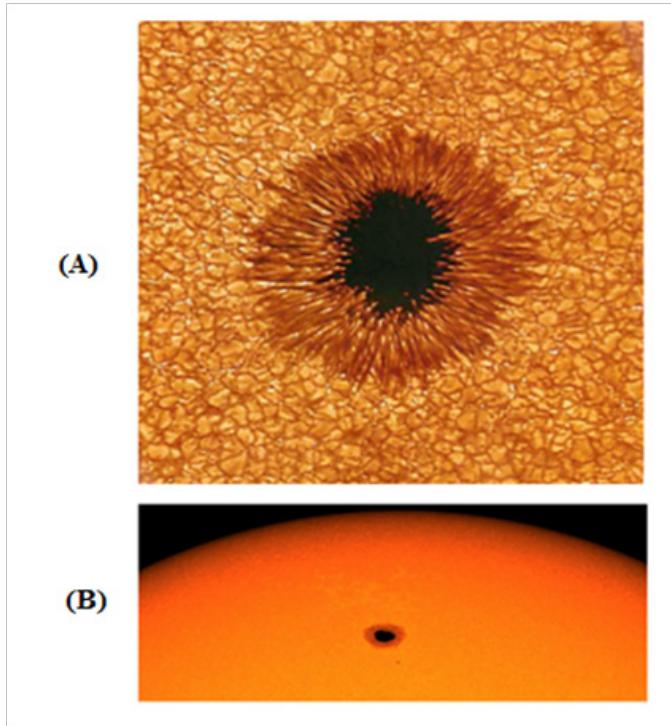


Figure 6 (A) An example of single spots (Courtesy of the Big Bear Observatory). (B) A remarkable example of a single spot without any other active feature in a large area of the solar disk (NASA's Dynamics Observatory). (A) and (B) are not the same spot.

It has also been considered that some irregularities in the convection flow tend to accumulate the magnetic fields at a few locations more than other places along the boundary. This feature is schematically illustrated in Figure 9 by combining two convective cell structures (in a negative unipolar region) constructed by Clark and Johnson²⁰.

In the left hand side of Figure 9, pores thus formed are assembled along the cell boundary to form small spots. Their assembly forms larger spots, as shown in the right hand side of Figure 9. This process is morphologically called '*coalescence*'.⁹ It is likely that the single spots in Figure 7A and 7B resulted from this process of coalescence. Parker²¹ suggested one possible way of the coalescence by assuming that small spots have vortex motions.

Thus, in a positive unipolar magnetic region in active areas, several positive pores might form a positive small single spot by their coalescence, and eventually a *positive single spot* by the coalescence of small spots as shown in Figure 7A & 7B.

Therefore, the fact that *positive monopole-like single spots* are produced by the coalescence within a *positive unipolar region* (vice versa) can satisfy the condition of $\text{div} \cdot \mathbf{B} = 0$, so long as unipolar magnetic regions are a product of the solar internal dynamo, as discussed in the above section.

In summary, the fact that a positive unipolar region is the birth place of a positive single spot (vice versa), *not elsewhere*, has advanced

toward our understanding of this single spot formation. Therefore, the problem (1) and observation (4) may be explained in this way.

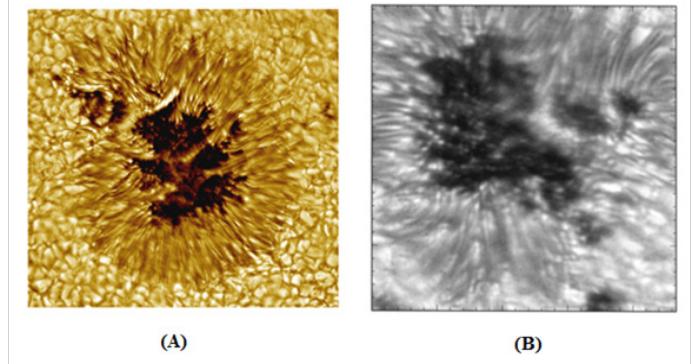


Figure 7 (A) and (B) are two high resolution images of a single spot, which show that a single spot consists of smaller spots, not a single column of magnetic flux; (a) Tritschler and Schmidt¹⁹ (b) NASA Image collection. .

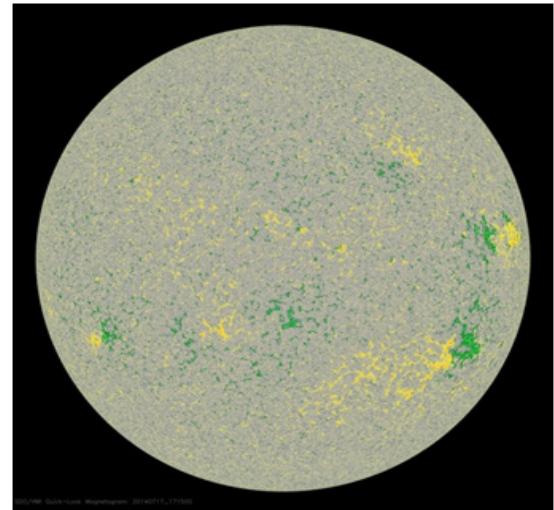


Figure 8 A high resolution image of the solar magnetic field (NASA).

This point will become more important in considering a pair of spots. In the next section, it will be shown that the above-mentioned process occurs at the boundary of a positive and a negative unipolar magnetic region of the opposite polarity, forming a pairs of spots.

Pair of spots

Various ideas of the formation of pairs of spots before Babcock⁸ was mentioned by Cowling⁸. After Babcock's paper, a large number of observations have been published by supporting the RF model.²² Although Getling²³ showed recently that their observations do not support the RF model, we still face the problems (2, 3, 5) unsolved today. Now, the problem (1, 4) may now find its possible explanation in Section 3. In this section, it is suggested that this explanation of the problem (1) may be quite important in considering the problems (2 and 3). Actually, the problem 5 gives us some hint in considering the remaining problems.

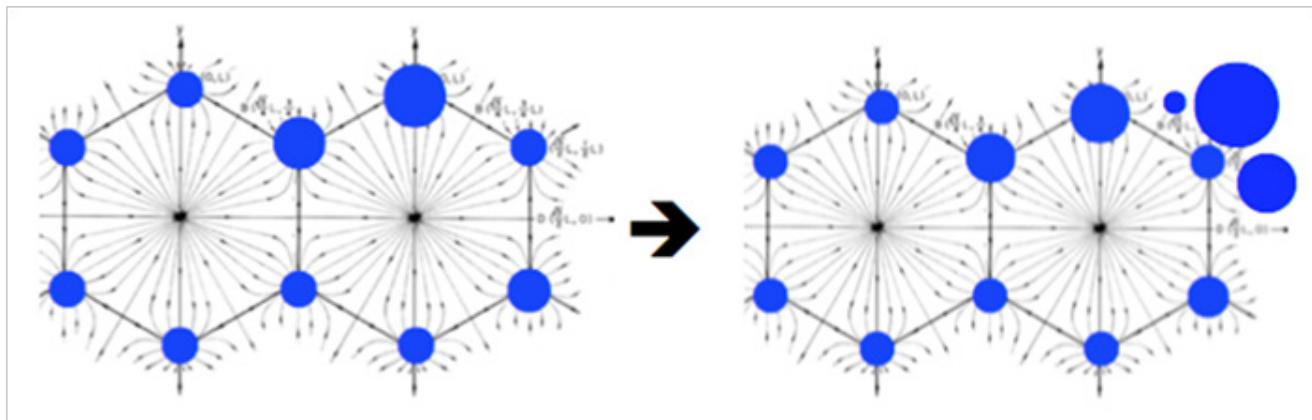


Figure 9 The figure illustrates schematically how pores, small spots and larger spots may be formed along the cell boundary. Left: A schematic illustration of the formation of pores by combining two convection cells in a negative unipolar region (blue), which is simulated by Clark and Johnson²⁰. Right: A schematic illustration of the formation of a single spot by the coalescence of pores and small spots.

Formation of a pair of spots

(a) Importance of the boundary of two unipolar regions

If magnetic buoyancy (or an upward flow) is the only reason for the rise of a magnetic flux tube in the RF model, a pair of positive and negative spots may appear together randomly at any location on the photosphere, even in the middle of unipolar regions. This is not the case. Thus, this is a serious problem for the RF model.

In considering the pair formation, we learned that *a sunspot pair appears across the boundary of positive and negative unipolar regions, and further positive one in a positive unipolar region (and vice versa)*⁹⁻¹¹. Thus, the boundary of two unipolar regions is a crucial fact in considering the pair formation; actually, this is collateral and consistent with the fact mentioned earlier, in that positive spots of the pair appear in a positive unipolar region (vice versa).

Thus, as we discuss later the delay of appearance of *f*-spots, there is a possibility that a positive *p*-spot in a positive unipolar magnetic region may induce a negative *f*-spot in a negative unipolar region across the boundary (vice versa), if a positive and a negative unipolar magnetic regions develop the connection by magnetic field lines.

McIntosh⁹ described morphologically the development of sunspot pairs in great details. He emphasized particularly the fact that a pair of sunspots develops along an elliptical cell-like structure of the size of $2-3 \times 10^4 \text{ km}$. Figure 10 is an example of the elliptical formation of sunspot pairs; note that as Figure 10 shows, larger spots consist of coalesced smaller spots. Many examples of the elliptical structure are shown in his paper, so that this elliptical formation is a typical way by which a spot pair develops. This is another fact which is not easy to explain by the RF model.

Figures 11 shows a higher resolution image around the boundary of a positive and negative unipolar regions.

Around the boundary of the positive and negative unipolar magnetic regions, one can see several U-shaped structures (lying horizontally), positive ones in a positive unipolar region (vice versa). At the boundary region, the convection flow appears to produce such magnetically split structures, namely positive and negative half-

cells in terms of the polarity of the magnetic field. Figure 10 shows schematically a combination of positive and negative U-shaped structures at the boundary.

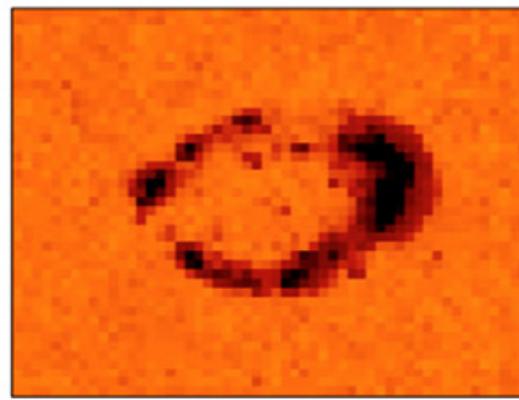


Figure 10 An example of a well developed pair of sunspots on October 27, 2003, showing the development of a pair of spots along an elliptical cell-like structure (Courtesy of the Kitt Peak Solar Observatory). In this spot pair, a large spot on the right side is negative, and spots on the left side are positive. It can be seen that the large negative spot may consist of several small spots. There is also a difference of the size between the two spots, which will be discussed later. For the blue and red U-shaped structures, see the text in the following.

If one of a positive U-shaped half-cell in the positive unipolar region at the boundary can induce the coalescence in a negative U-shaped structure, which is connected by magnetic field lines, two U-shaped structures of the opposite sign can form a pair of spots along the elliptical boundary. Figure 12 shows schematically how positive and negative U-shaped cells at the boundary form a pair of spots by

combining the cell structure simulated by Clark & Johnson²⁰. This is one way by which McIntosh's elliptical structure of a pair of spots can be explained. This inducing process will be discussed further in the following in connection with problems (2,3).. It may be the reason why a pair of spots tends to appear at the boundary of two unipolar magnetic region of the opposite polarity.

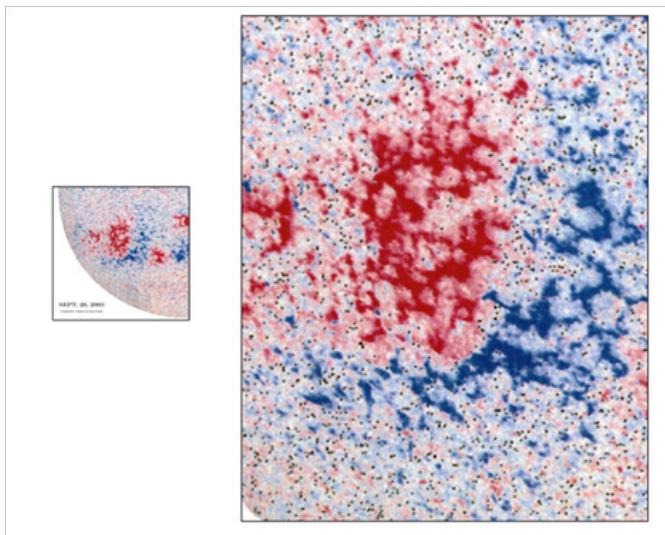


Figure 11 They show the magnetic network structure near the boundary of two unipolar magnetic regions, September 28, 2003 (Courtesy of the Kitts Peak Observatory). It is crucial to note that there are positive (red) and negative (blue) U-shaped structures near the boundary.

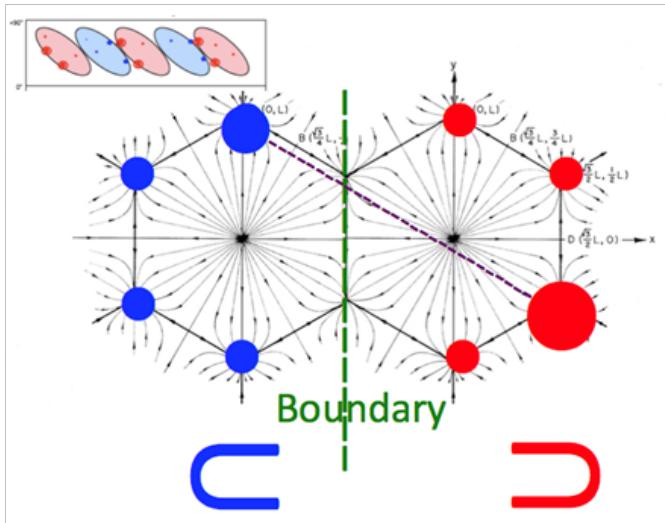


Figure 12 Schematic illustration of the magnetic network near the boundary of two unipolar magnetic regions, showing a pair of negative (blue) and positive (positive) U-shaped structures. It is likely that McIntosh's elliptical formation of a pair of spots may develop by connecting two U-shaped structures of the opposite polarity. One magnetic field line is shown to connect a large positive and negative spots. Upper left: A portion of Figure 5F. Middle: The convective cells at the boundary. Positive and negative cells are combined at the boundary. The network structure at the boundary is thus different from that shown in Figure 9. Bottom: Two U-shaped structure shown in Figure 10.

Positive pores and small spots may coalesce *along the positive U-shaped structure (which is a half of the boundary of a cell)* to form a small positive single spots and further eventually a larger single spot

(or spots) and vice versa. Eventually, spots in a positive U-shaped structure coalesce at its end and spots in a negative U-shaped structure coalesce at the end, forming a typical pair of spots, which consists simply of a pair of positive and negative spots. Figure 13 shows this transition and development. A number of such examples are shown by McIntosh⁹. Note that as mentioned earlier, both spots consist of many smaller spots.

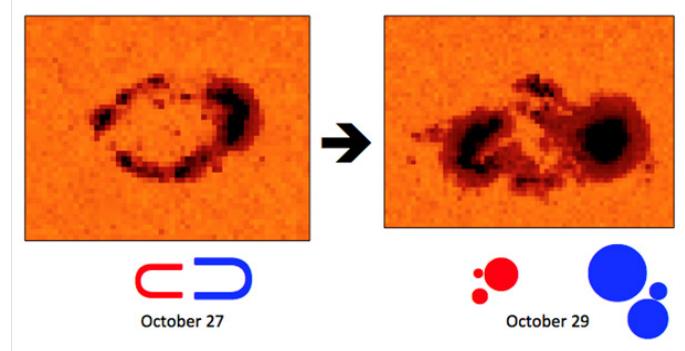


Figure 13 Transition from an elliptical formation of small spots to a simplified typical pair of sunspots, which are actually consisted of several smaller spots. The elliptical formation is produced by combination of two U-shaped formation.

Delay of *f*-spots

According to the RF model, a *p*-spot and *f*-spot should appear simultaneously. However, it has long been known that a pair of spots does not appear at the same time; as so named, *p* (primary) spots appear earlier than *f* (following) spots.^{3,4} There are only a few cases, in which *f*-spots appear first, so that such cases are rare.⁹

Thus, this delay of *f*-spot is indeed an important fact in considering the pair formation. In Figure 10, *p*-spots (the right hand side) appeared first, and this pair developed in a period of less than 5 hours as a *very faint* elliptical form; because of its faintness, the elliptical structure (the cell boundary) may not be easily recognized, so that it may be seen only as two (*p* and *f*) spots.

In the past, on the basis of the RF model, various attempts (such as the look-angle of the tube or branching of a tube below the photospheric surface) were made to explain the delay of the appearance of *f*-spots after *p*-spots, but those ideas cannot be confirmed by observations below the photospheric surface. The present way of explaining single spots and pairs of spots *based on the observed features* provides an opportunity to consider this problem which is contradictory to the RF model (thus, the problem cannot be explained without hypothesizing processes which cannot be proven at this time).

This inducing effect is possible, if some portions of positive and negative unipolar regions are connected by magnetic field lines. The formation of a *p*-side of the U-shaped half-cell could be communicated to *f*-side of U-shaped half-cell, because magnetic field lines are expected to be nearly equi-potential in the solar atmosphere, so that any electromagnetic changes associated with the coalescence (e.g. a converging flow, $V \times B$) in a *p*-side U will be communicated to a *f*-side U by Alfvén waves, inducing a similar coalescence at the other end of the magnetic field lines, namely the formation of *f*-spots of the opposite polarity.

Thus, some delay might be expected as the communication time. This might be a possible explanation on the problem (2). In

this respect, it may be noted that Kotov²⁴ showed that there is 10^{12} A of electric currents around a single spot, since it can be caused by the $V \times B$ process (Figure 14). It is in this way that the problem (5) provides an opportunity to discuss the delay, because the RF model does not allow the delay.

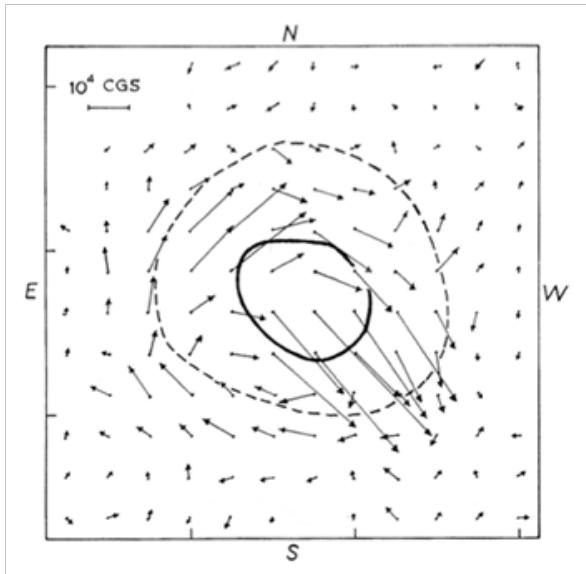


Figure 14 The electric current distribution around a single spot.²⁴

Unbalance of the magnetic flux

It is well known that there is often a significant unbalance of the magnetic flux between two spots of a pair.^{4,6} Figure 10 is a good such example. This unbalance is not easy to explain in terms of a simple RF model, although in the past some efforts were made by considering a complicated branching of the tube, which cannot be observed deep in the photosphere. However, so long as we consider a single magnetic flux tube (the RF model), the unbalance is a contradictory problem for the RF model, namely the problem (3).

When a *p*-spot is forming in a positive unipolar region, it does not have to induce an *identical f*-spot just across its neighboring unipolar region, because the field distribution in the two unipolar magnetic regions across the boundary or in the surrounding regions is not a mirror image of each other and thus different (Figure 11).

In this respect, as shown in Figure 15, X-ray observations of the corona show that some of the magnetic field lines in a pair of spots are connected to at least another pair of spots, even across the magnetic equator,²⁵ so that it is not necessary to consider the induction within just one pair of spots; another good example is seen in his Figure 5 of Aschwanden²⁶. The unbalance of the magnetic flux is contradictory to the RF model, so that it is not possible to explain the problem 3, but the present idea of the formation of a pair of spots provides an opportunity to consider the problem 3 on the basis of the problem 5.

The first appearance of *p*-spot

Actually, instead of considering the *delay* of *f*-spot, it is more appropriate to consider *why a p-spot appears first before a f-spot*. At this point, we cannot find definitive morphological observations in explaining why a *p*-spot appears first at the specific boundary of two unipolar regions, not the other side, except that the specific side of the

unipolar region tends to become more active first. In this respect, it may be noted that spots are surrounded by faculae, indicating specific activity around single sunspots.¹ Svalgaard and Wilcox²⁷ noted this particular boundary and called it the "Hale boundary". This may be related to the structure of unipolar regions (Figure 3C).

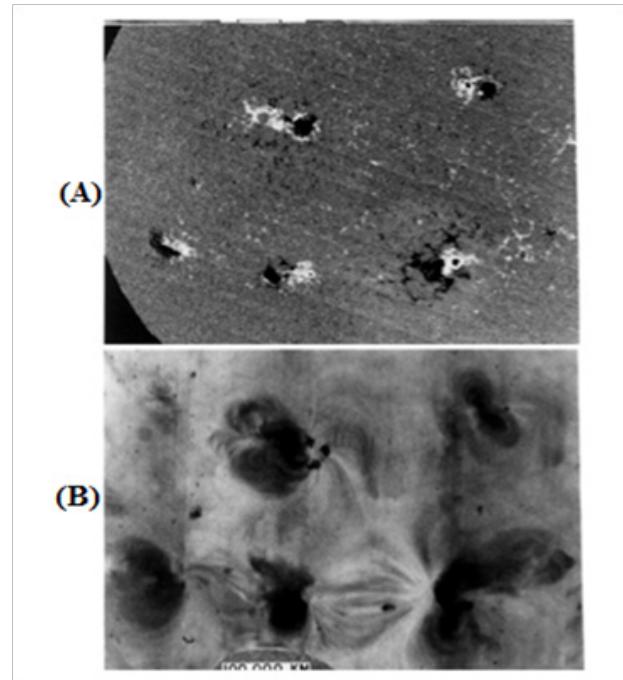


Figure 15 (A) The photospheric magnetic field distribution, in which five sunspot pairs are seen; the magnetic equator crosses the middle of the record horizontally. (B) The corresponding X-ray image of the corona. Some spot pairs are connected each other even across the magnetic equator.²⁵

The sunspot cycle

Any idea of spot formation must explain the equatorward shift during the sunspot cycle. There are a number of dynamo theories on the Butterfly diagram, explaining the equatorward shift of the formation of sunspot pairs during the sunspot cycle. Since we are considering only *morphologically* the formation of sunspots in this paper, it may be noted that there is an interesting observation which shows that an east–west belt of 'torsional oscillation' (rotation or anti-parallel flow) on the photosphere shifts from poleward to equatorward during each sunspot cycle;²⁸ Figure 16. McIntosh⁹ showed that there is a counter flow along such a belt and that solar activities are high certain locations along the belt. How such a dynamical belt can stimulate the formation or coalescence of spots at the specific boundary of two unipolar regions is beyond the scope of this paper.

In this respect, McIntosh⁹ noted that sunspots, which tend to show the Coriolis effect, exhibit drastic spot coalescence. Since the rotation speed of the sun at the equator is about 2 km/s and plasma velocity around sunspots is only about several km/s, one does not expect much the Coriolis effect on the photosphere. However, a rare example of vertical sunspot group was reported by Akasofu²⁹; Figure 17 shows both the vertical sunspot group and a nearby single spot. However, dynamical processes and their relation to the formation of sunspots are beyond the scope of this paper, except to mention that Parker²¹ considered vortex motions of small spots in explaining the coalescence.

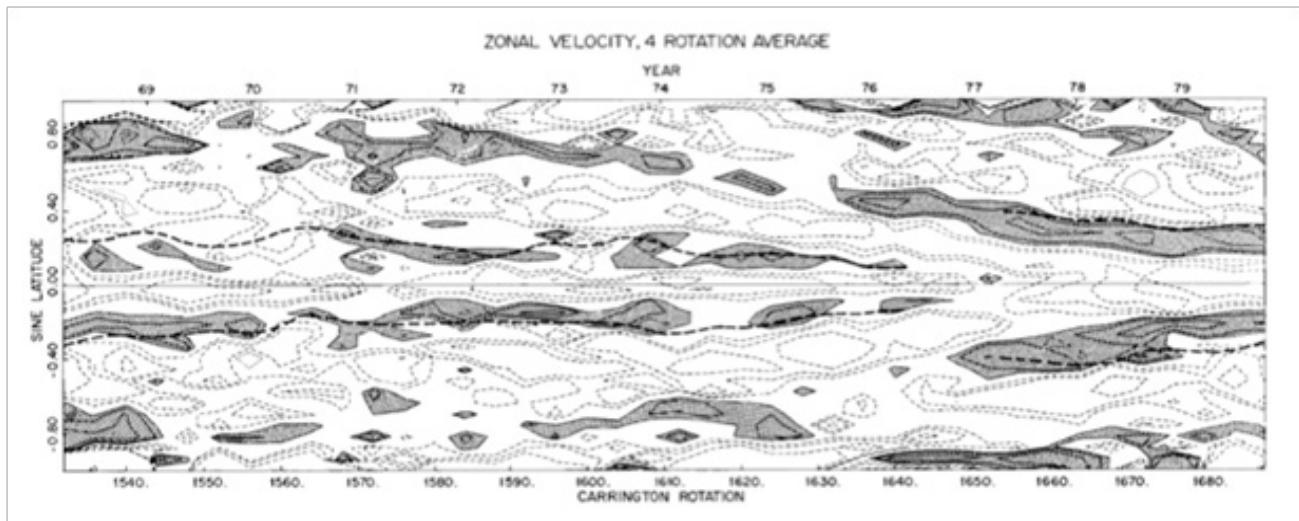


Figure 16 The equatorward shift of the east–west belt of torsional oscillation during a few solar cycles.²⁸

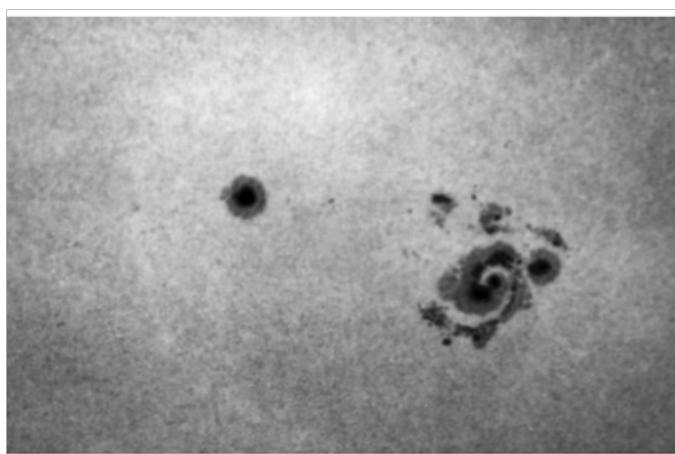


Figure 17 Vortical distribution of a sunspot group observed on February 19, 1982 (Courtesy of W. Livingston and F. F. Recely at the Kitt Peak Solar Observatory), together with a nearby single spot.

Conclusion

Although the RF model has long been well accepted, and there are a large number of supporting observations,²² the three long-standing problems (1–3) and two new observations (4) and (5) seem to be difficult to explain, so long as we assume a rising magnetic flux tube. The present study began as an attempt to study the three (1–3) problems, but it is found to be very difficult to proceed on the basis of a simple RF model, since they are obviously contradictory to the RF model. Fortunately, the newer problems (4, 5) provided a hint in considering the problem (1, 2, 3) and further all of them together.

This paper is the full paper of a paper published earlier as a short report.³⁰ It is hoped that this paper will initiate even more new ideas in considering the formation of sunspots in the future. It is suggested that the following points should further confirmed morphologically and theoretically, regardless of any theory.

1. Nature and origin of unipolar magnetic regions.

2. Process of the coalescence of pores to spots.
3. Induction process of a *f*-spot by a *p*-spot.
4. Unbalance of the magnetic flux between a *p*-spot and a *f*-spot.
5. Appearance of *p*-spots before *f*-spots.
6. Nature of the Hale boundary.
7. Relation between Howard–LaBonte’s observation and the spot formation.

Acknowledgements

The author would like to acknowledge Dr. Kees de Jager for his comments on drafts of this paper. He would like to acknowledge also the Kitt Peak Solar Observatory, the Big Bear Solar Observatory and the NASA Image Collection for their magnetic records and spot images used in this paper.

Conflict of interest

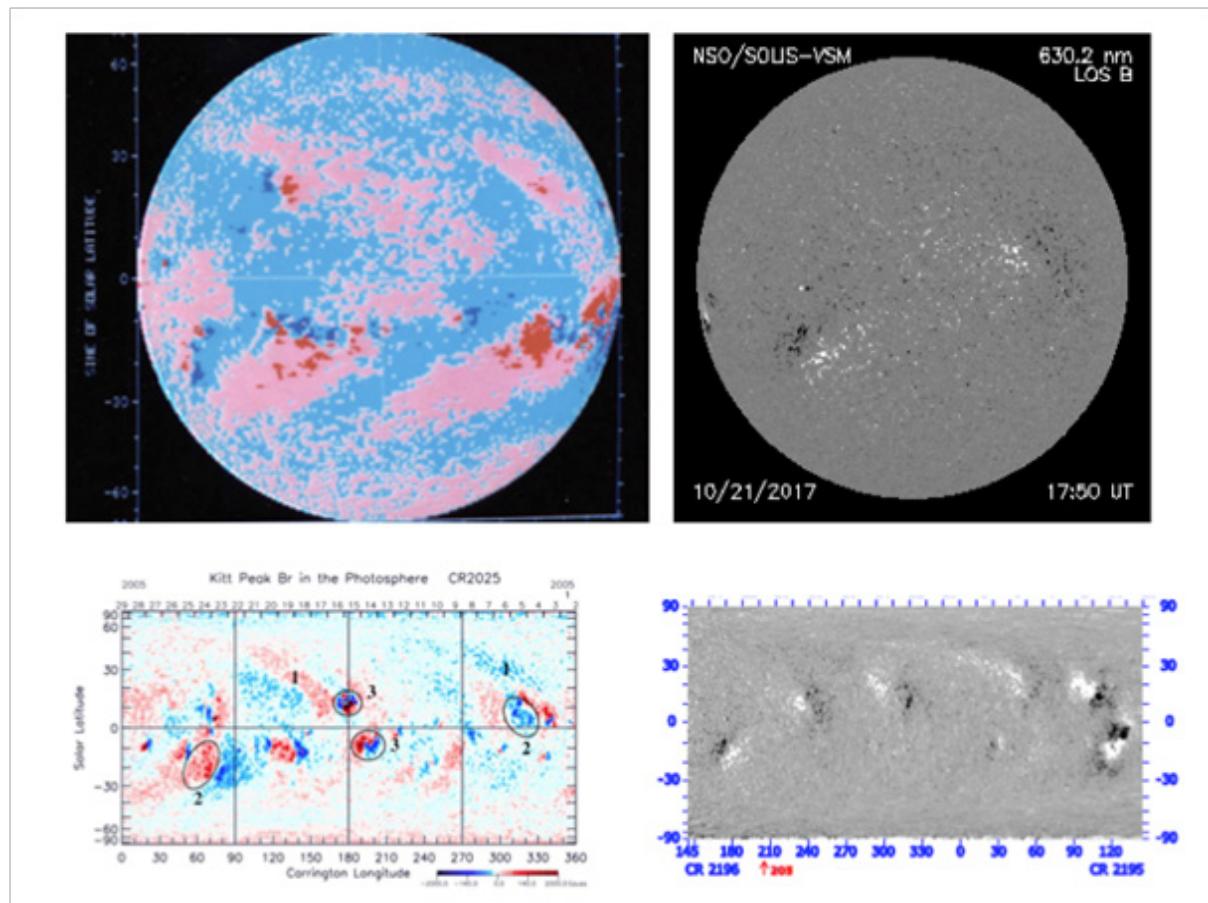
Author declares that there is no conflict of interest.

References

1. Abetti G. *The Sun*, Faber and Faber. UK; 24 Russel Square; 1951.
2. Cowling TG. *Solar Electrodynamics*. In: Kuiper GP, editor. USA: The University of Chicago Press; 1953. p. 532–591.
3. Kiepenheuer KO. *Solar activity*. In: Kuiper GP, editor. USA: The University of Chicago Press; 1953. p. 322–463.
4. De Jager C. *Structure and dynamics of the solar atmosphere*. In: Flugge S, editor. Germany: Springer-Verlag; 1959. p. 80–362.
5. Menzel DH. *The Sun*. Canada: The Blakiston & Company; 1959.
6. Bray RJ, Loughhead RE. *Sunspots*. USA: Dover Publications; 1964.
7. Solanki SK. Sunspots:Overview. *The Astronomy and Astrophysics Review*. 2003;11(2–3):153–286.
8. Babcock HW. The topology of the sun’s magnetic field and the 22–year cycle. *The Astrophysical Journal*. 1961;133:B572–B587.

9. McIntosh PS. *The Physics of Sunspots*. In: Crom LE, Thomas JH, editors. New Mexico: Sacramento Peak Observatrор; 1981. p. 7–54.
10. Akasofu SI. Single spots, unipolar magnetic regions, and pairs of spots. *Geophysical Research Letters*. 2014;41(11):3698–3700.
11. Akasofu SI. Single spots, unipolar magnetic regions, and pairs of spots: 2 the development of sunspot pairs and the Hale boundary. *Geophysical Research Letters*. 2015;42(8):2571–2576.
12. Leighton RB. A magneto-kinematic model of the solar cycle. *The Astrophysical Journal*. 1969;156:1–26.
13. Hathaway DH. The solar cycle. *Living Reviews in Solar Physics*. 2010;7:1.
14. De Jager C, Akasofu SI, Duhau S, et al. The remarkable recent phase transition in the solar dynamo. *Space Science Reviews*. 2016;201(1–4):109–145.
15. McIntosh PS, Wilson PR. A new model for flux emetgence and the evolution of sunspots and the large-scale fields. *Solar Physics*. 1985;97(1):59–79.
16. Wilson PR. The reversal of the solar polar magnetic fields III: The large scale fields and the first major active regions of cycle 22. *Solar Physics*. 1992;138(1):11–21.
17. Nakagawa Y. *A numerical study of the solar cycle, Solar Magnetic Fields*. In: Howard R, editor. Netherlands: D Reidel Publishing Company; 1971.
18. Noyes RW. *The Sun, Our star*. USA: Harvard University Press; 1982: 263 p.
19. Tritschler A, Schmidt W. Sunspot photometry with phase diversity I: Methods and global sunspot parameters. *Astronomy & Astrophysics*. 2002;382(3):1093–1105.
20. Clark A, Johnson AC. Magnetic field accumulation in supergranules. *Solar Physics*. 1967;2(4):433–440.
21. Parker EN. Vortex attraction and the formation of sunspots. *The Astrophysical Journal*. 1992;390(1):290–296.
22. Itonidis S, Zhao J, Kosovichev A. Detection of emerging sunspot regions in the solar interior. *Science*. 2011;333(6045):993–996.
23. Getling AV, Ishikawas R, Buchnev AA. Development of active regions: Flows, magnetic-field patterns and bordering effect. *Solar Physics*. 2016;291(2):371–382.
24. Kotov VA. *Solar Magnetic Fields*. In: Howard R, editor. Netherlands: D Reidel Publishing Company; 1971. p. 213–219.
25. Sheeley NR. Energy released by the interaction of coronal magnetic fields. *Solar Physics*. 1976;47(1):173–180.
26. Aschwanden M. *Physics of the Solar Corona*. Germany: Springer; 2005.
27. Svalgaard L, Wilcox JM. The Hale solar sector boundary. *Solar Physics*. 1976;49(1):177–185.
28. Howard R, La Bonte BJ. The sun is observed to be a torsional oscillator with a period of 11 years. *The Astrophysical Journal*. 1980;239:L33–L36.
29. Akasofu SI. Vortical distribution of sunspots. *Planetary and Space Science*. 1985;33(3):275–277.
30. Akasofu SI. A new morphological theory on the formation of sunspots. *Physics & Astronomy International Journal*. 2018;2(3):163–167.

Appendix



Appendix: Comparison between the Kitt Peak data (left) and recent NASA data (right). It can be seen that the continuous color presentation is more suited than the black–white presentation in the present work.